

PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

Registration accuracy between whole slide images and glass slides in eeDAP workflow

Qi Gong, Benjamin P. Berman, Marios A. Gavrielides, Brandon D. Gallas

Qi Gong, Benjamin P. Berman, Marios A. Gavrielides, Brandon D. Gallas, "Registration accuracy between whole slide images and glass slides in eeDAP workflow," Proc. SPIE 10581, Medical Imaging 2018: Digital Pathology, 1058118 (6 March 2018); doi: 10.1117/12.2293189

SPIE.

Event: SPIE Medical Imaging, 2018, Houston, Texas, United States

Registration accuracy between whole slide images and glass slides in eeDAP workflow

Qi Gong¹, Benjamin P. Berman², Marios A. Gavrielides¹, Brandon D. Gallas¹

¹Division of Imaging, Diagnostics, and Software Reliability, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland

²Division of Radiological Health, Office of In Vitro Diagnostics and Radiological Health, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland

ABSTRACT

The purpose of this study is evaluating registration accuracy of evaluation environment of Digital and Analog Pathology (eeDAP). eeDAP was developed to help conduct studies in which pathologists view and evaluate the same fields of view (FOVs), cells, or features in a glass slide on a microscope and in a whole slide image (WSI) on a digital display by registering the two domains. Registration happens at the beginning of a study (global registration) and during a study (local registration). The global registration is interactive and defines the correspondence between the WSI and stage coordinates. The local registration ensures the pathologist evaluates the correct FOVs, cells, and features. All registrations are based on image-based normalized cross correlation. This study evaluates the registration accuracy achieved throughout a study. To measure the registration accuracy, we used an eyepiece ruler reticle to measure the shift distance between the center of the eyepiece and a target feature expected in the center. Two readers independently registered 60 FOVs from 6 glass slides, which covered different tissue types, stains, and magnifications. The results show that when the camera image is in focus, the registration was within 5 micrometers in more than 95% of the FOVs. The tissue type, stain, magnification, or reader did not appear to impact local registration accuracy. The registration error was mainly dependent on the microscope being in focus, the scan quality, and the FOV content (unique high-contrast structures are better than content that is homogeneous or low contrast).

1. INTRODUCTION

Digital pathology (DP) incorporates the acquisition, management, and interpretation of pathology information generated from pathology images or from digital scans of the whole area of glass slides referred to as whole slide images (WSI). The potential benefits of DP include telepathology, digital consultation and slide sharing, pathology education, indexing and retrieval of cases, and the use of automated image analysis [1-5].

The evaluation environment of Digital and Analog Pathology (eeDAP) [1] was developed to conduct studies in which pathologists can evaluate the same fields of view (FOVs), cells, or features in a glass slide on a microscope and in a WSI on a digital display. eeDAP studies can be used to evaluate the performance of pathologists interpreting WSI images with the performance on the microscope as the baseline for comparison. eeDAP studies can also be used to

generate microscope-based annotations on the WSI to train and evaluate image analysis programs. These use cases are achieved by registering the glass slide to the WSI enabled by eeDAP. There are two parts of registration: global registration and local registration. Both of them are based on image-based cross correlation [6] between the WSI and the camera image of the microscope FOV. The WSI is large (up to 100k by 120k pixels under 40X [7]). It consumes significant amount of time to manually correlate a WSI location with the location on the glass slide under a microscope.

The goal of this study is to measure the registrations accuracy of eeDAP. An eyepiece ruler reticle is used to measure the shift distance between the center of eyepiece and a target feature expected in the center. Two readers independently registered 60 FOVs from 6 glass slides.

2. MATERIALS AND METHODS

eeDAP includes hardware components (computer, microscope, eyepiece camera, and microscope step motor and stage) and software to control the stage, perform registration, and collect data [1]. Figure 1 shows the eeDAP workflow. Registration is performed at the beginning of a study (global registration) and during a study (local registration). The global registration is interactive and defines the relationship between the WSI and stage coordinates. Local registration ensures that the pathologist evaluates the correct FOVs. Local registration can happen automatically or at the request of the user, and there are two types that we describe below: with and without padding. Focusing the microscope is important to local registration. Focusing the microscope is not automated in this work; it is the job of the pathologist. It could be automated in the future.

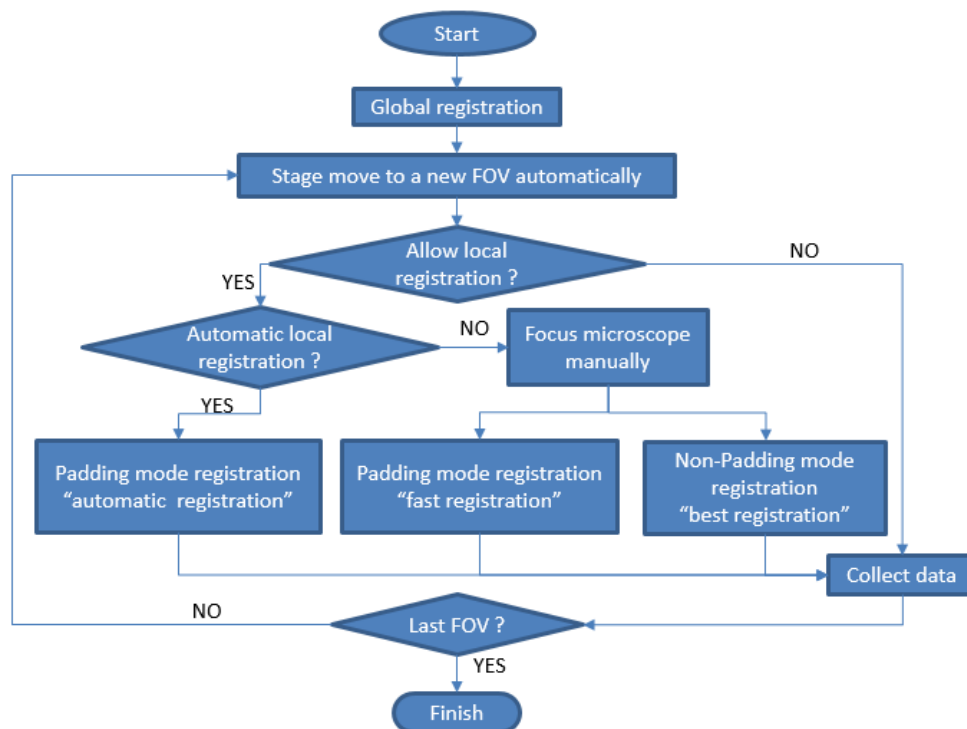


Figure 1. eeDAP workflow

All image registrations are based on normalized 2D cross correlation [6] and use a patch of the WSI and some or all of the camera image of the microscope FOV. The full camera image of the microscope FOV typically does not cover the eyepiece FOV, and we can use a demagnification camera adaptor to mount the camera on the microscope to improve the FOV coverage. The camera image is rescaled to the scale of the WSI FOV to implement the cross correlation. The registration by cross correlation convolves the two images after a color-to-grayscale transform [8]. The convolution result is high where the images are similar. Thus, the direction and distance from the peak of the

convolution result to its center is the shift between the camera image and the WSIFOV. More details can be found in Gallas et al [1].

There are two modes of registration: Padding mode and Non-Padding mode. In Padding mode, padding the base image (the larger of the two images being compared) with zeros allows cross correlation to find the target (the smaller of the two images being compared) when it is on the boundary of the base image. However, the padding can lead to registration errors (finding the wrong location). In Non-Padding mode, the target image must be completely contained within the base image for it to be found. This mode has higher registration accuracy for the center image area, but it cannot find the target if it is located on the boundary of the base image. The global registration is based on Padding mode. Local registration can be based on either mode. When local registration is done automatically, it uses Padding mode.

Table 1 Study Hardware

Microscope	Camera	Microscope step motor	Microscope stage
Axioplan 2 Imaging	Point Grey Flea2 Color	Ludl MAC5000	Marzhauser SCAN 8Preparate Axioplan2

Table 2: Study slides and scanner details



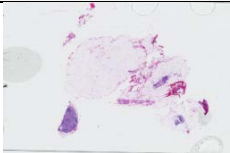
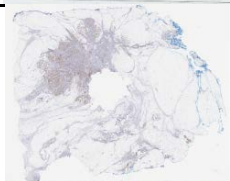
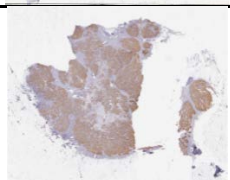
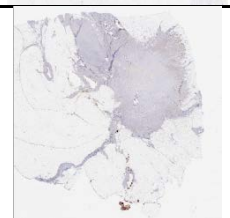
Tissues	Stain	WSI information	Thumbnail
Canine oral melanoma	H&E	Magnification: 40X Spatial Resolution: 0.25 μ m Size: 91,391 x 17,808 pixels Size: 22.85 mm x 4.45 mm	
Canine oral melanoma	pHH3	Magnification: 40X Spatial Resolution: 0.25 μ m Size: 95,199 x 19,086 pixels Size: 23.80 mm x 4.77 mm	
Human lymph node	H&E	Magnification: 40X Spatial Resolution: 0.23 μ m Size: 163,840 x 107,520 pixels Size: 37.68 mm x 24.73 mm	
Human breast	ER	Magnification: 20X Spatial Resolution: 0.50 μ m Size: 56,015 x 45,009 pixels Size: 28.01 mm x 22.50 mm	
Human breast	HER2 (+3)	Magnification: 20X Spatial Resolution: 0.50 μ m Size: 34,198 x 27,840 pixels Size: 17.10 mm x 13.92 mm	
Human breast	HER2 (+1)	Magnification: 20X Spatial Resolution: 0.50 μ m Size: 30,598 x 31,674 pixels Size: 15.30 mm x 15.84 mm	

Table 1 shows the hardware that was used in the study and Table 2 shows and describes the glass slides. The study covers different tissue types, stains, and magnifications in order to evaluate registration accuracy over these factors. These factors may affect glass slide and WSI qualities [9, 10], and may also affect eeDAP performance.

10 FOVs were chosen per slide. All FOVs had a small identifiable target in the center. A virtual reticle was used to locate the target in the eeDAP view of the WSI (Figure 2-A). Some of the FOVs were purposefully selected to stress-test the system; many had homogeneous, sparse, or low contrast content. The readers looked through the microscope and used a rotatable ruler eyepiece reticle (Figure 2-B) to measure the distance between the target and the center of the eyepiece view. The ruler has 100 divisions. Each division is 0.1mm (2.5 μm at 40X, 5.0 μm at 20X).

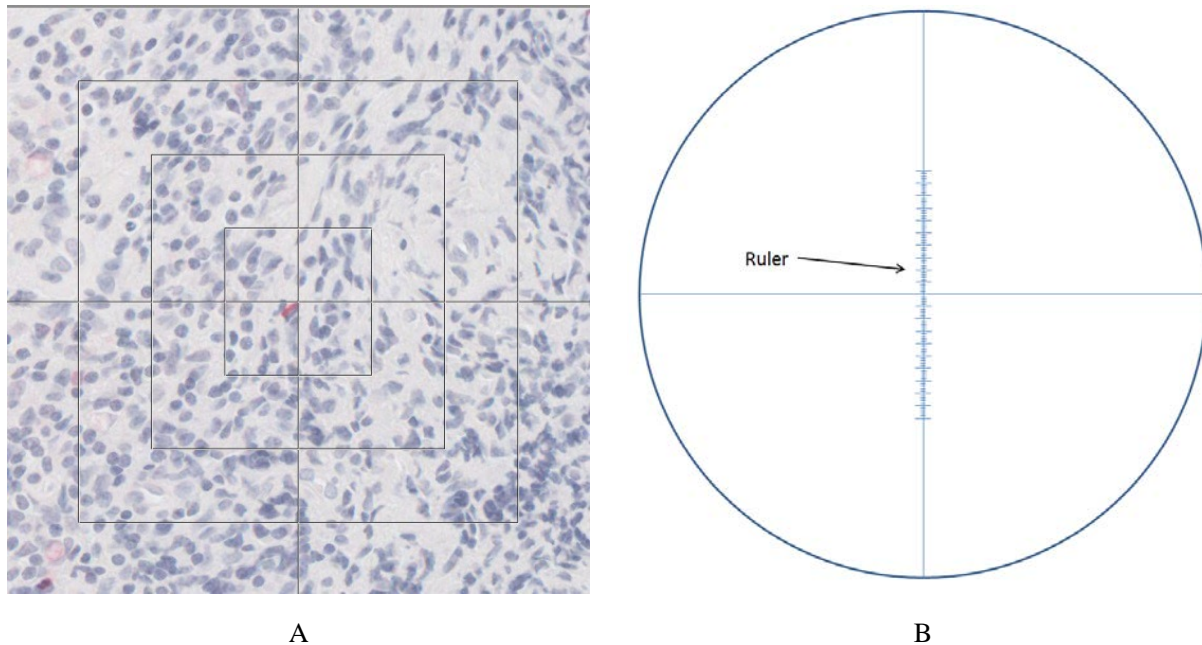


Figure 2: A) Virtual reticle indicates the target in the eeDAP view of the WSI (top-right corner of a red cell). B) Eyepiece ruler reticle.

There were two parts in this study: Global Registration Study and Local Registration Study. The Global Registration Study evaluated global registration and stage motion. In the Global Registration Study, the reader performed global registration once and measured the error for each FOV one by one. During this study, the reader did not perform any local registrations and measured one registration error per FOV. In the Local Registration Study, the reader began with global registration. During the study, for each FOV, the reader used local registrations to refine the global registration, and then measured the errors. As illustrated in Figure 1, there are three types of local registrations: 1) Registration with padding before focusing the microscope: this type uses the previous FOV focus plane and is done automatically to speed up system; 2) Registration with padding after manually focusing the microscope; and 3) Registration without padding after manually focusing the microscope.

Registration error can also be impacted by the distance traveled by the stage and differences in the focusing planes of different FOVs (within and across slides). Therefore, we repeated the study for the same FOVs but with two different orders: 1) List order: the reader measured all FOVs on one slide and then went to the next slide; 2) Random order: The FOVs appeared in a random order, with sequential FOVs potentially occurring on different slides. The travel distances and the focus differences between FOVs was larger in the Random order study compared to the List order study. In total, each reader measured eight registration errors for all 60 FOVs: (List order vs. Random order) \times (no local registration + 3 local registration types).

3. RESULTS

3.1 Global Registration Study

Figure 3 shows the histogram of registration errors for two readers measured in the Global Registration Study. The means of measured errors are $33.14 \mu\text{m}$ ($SD = 23.39$) and $42.02 \mu\text{m}$ ($SD = 32.40$) for the two readers, respectively. The diameter of a cell measured in this study is about $15 \mu\text{m}$ [11, 12]. This implies that global registration can find large features but cannot point to a cell. Results show wide variability in terms of registration error supporting the need for the local registration step.

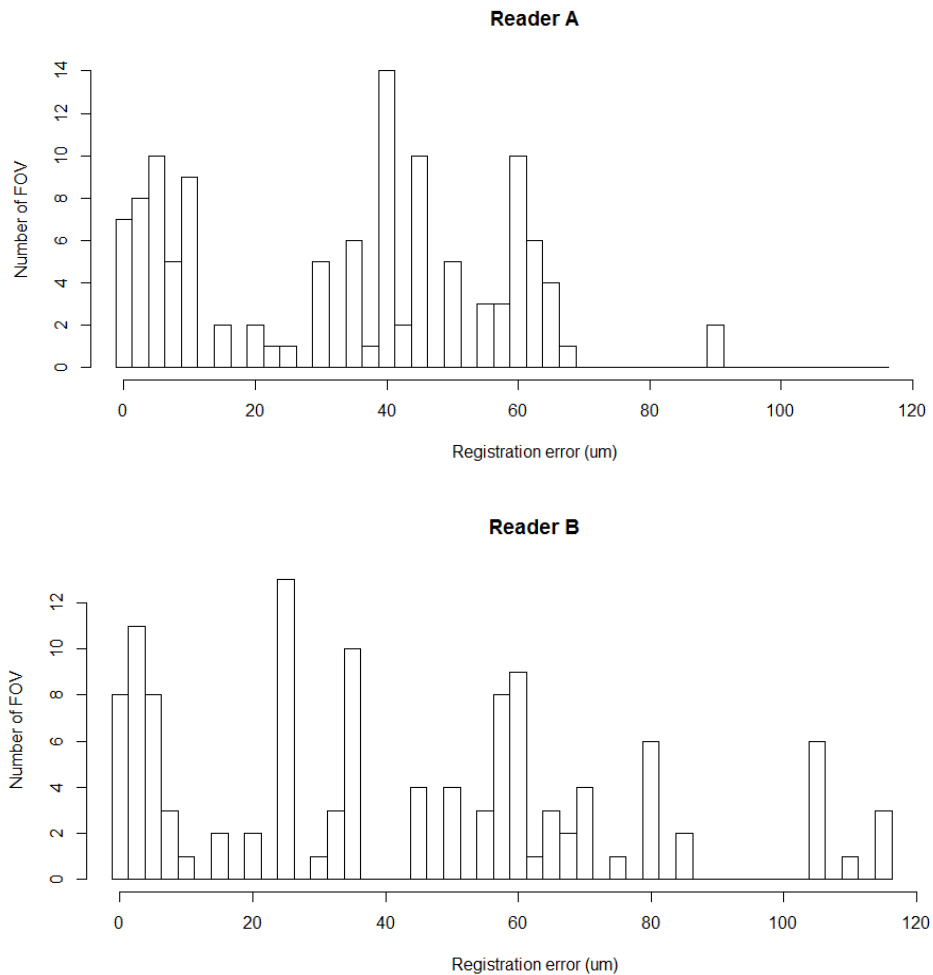


Figure 3. Errors for two readers from the Global Registration Study

3.2 Local Registration Study

Figure 4 shows the histograms of registration errors of the three local registrations. The errors are either smaller than or equal to $5 \mu\text{m}$ or considerably larger ($> 100 \mu\text{m}$), where $2.5 \mu\text{m}$ is the smallest measurable error at 40X and $5 \mu\text{m}$ is the smallest measurable error at 20X. The majority of the observations were $5 \mu\text{m}$ or less. Given this result and the

fact that the diameter of a cell is about $15\ \mu\text{m}$, we shall bin the registration error into two bins: errors within $5\ \mu\text{m}$ and errors greater than $5\ \mu\text{m}$. These findings also show the effectiveness of local registration.

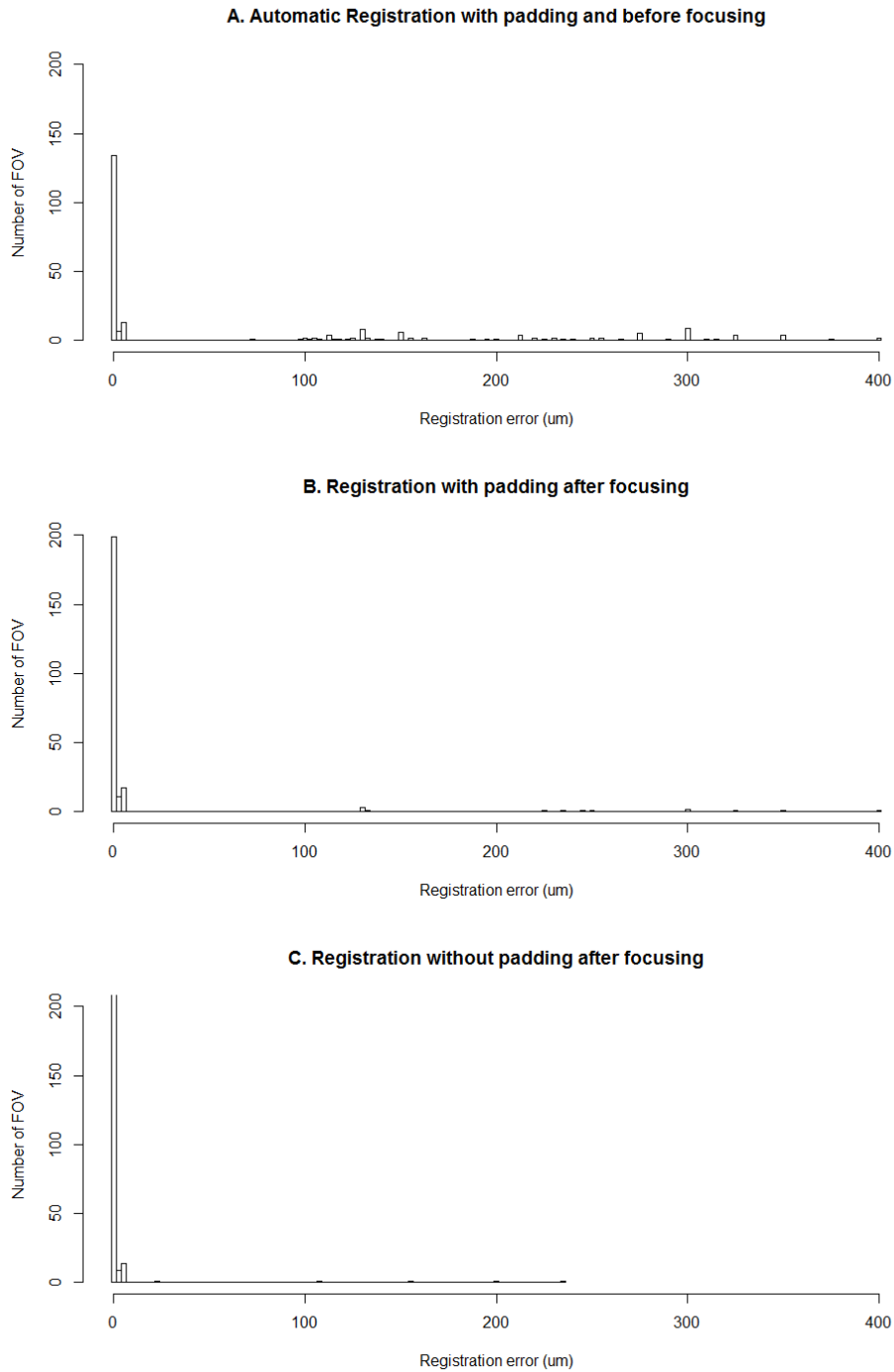


Figure 4. Three local registrations errors for two readers from Local Registration Study

Automatic Registration with Padding and Before Focusing

Table 3 shows the Padding Registration (without focusing) errors from two readers measured in Local Registration Study. Based on two readers' results, 76.7% FOVs obtained registration error within 5 μm in list order studies, and 51.7% FOVs obtained small registration error within 5 μm in random order studies. For both readers, the list order studies outperformed the random order studies on all slides. The reason for this is that because the system uses previous FOV focus plane do registration, each FOV is more likely to be in focus during the list-mode study since the order of FOVs is grouped by slide. Since different slides have different thicknesses, the focus planes of different slides are likely different. Figure 5 shows images (after color-to-grey operation) for an FOV that achieved a small registration error ($\leq 5 \mu\text{m}$) in the list order study but a large registration error ($>5 \mu\text{m}$) in the random order study. Random order study before focusing the camera image (Figure 5-B) is more blurred than list order study before focusing the camera image (Figure 5-C) and the registration error of 5-B is significantly larger than that of 5-C. It is worth noting that 5-C is not a perfectly focused image either.

Table 3: Accuracy of Registration with Padding Before Focusing (fraction of FOVs with error $\leq 5 \mu\text{m}$)

WSI name	List order		Random order	
	Reader A	Reader B	Reader A	Reader B
Human breast HER2 (+1)	6/10	6/10	4/10	4/10
Human breast HER2 (+3)	9/10	9/10	8/10	7/10
Human breast ER	7/10	6/10	6/10	7/10
Human lymph node H&E	6/10	7/10	3/10	4/10
Canine oral melanoma H&E	9/10	9/10	3/10	6/10
Canine oral melanoma pHH3	9/10	9/10	4/10	6/10
pool over images	46/60 (76.7%) SE = 0.05	46/60 (76.7%) SE = 0.05	28/60 (46.7%) SE = 0.06	34/60 (56.7%) SE = 0.06
pool over readers	92/120 (76.7%)		62/120 (51.7%)	

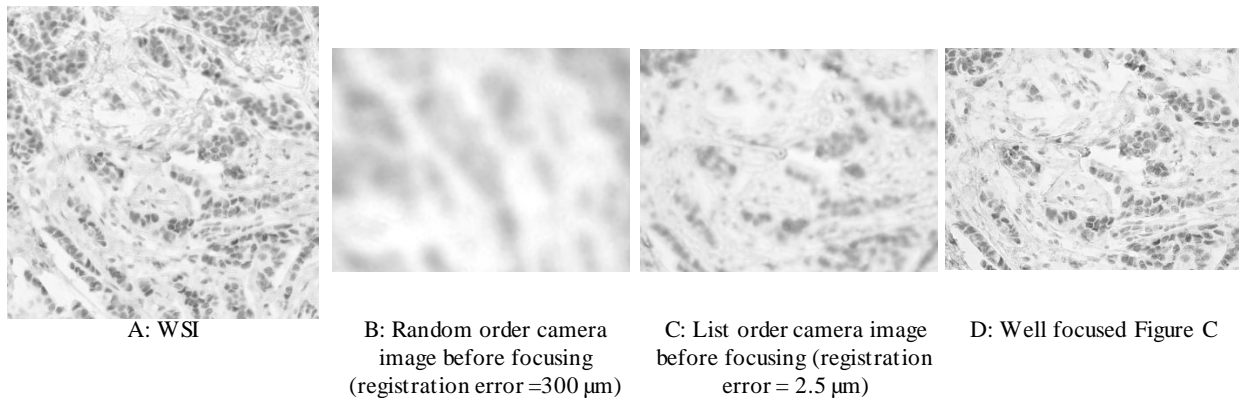


Figure 5: An example FOV (after color-to-grey operation) that has a small error ($\leq 5 \mu\text{m}$) in list order study and a large error ($>5 \mu\text{m}$) in random order study

Registration With and Without Padding after Focusing

Table 4 shows the accuracy from registration with and without padding after focusing the microscope. In these studies, the reader focused the microscope for all FOVs before registration. This essentially eliminates the impact of study order. For registration with padding, 94.2% of the time the registration was within 5 μm in list order studies and 95.0%

in random order studies, respectively. For registration without padding, the percentages are 96.7% and 99.2% for list order studies and random order studies.

Table 4: Accuracy of registration with and without padding after focusing the microscope (fraction of FOVs with error $\leq 5 \mu\text{m}$)

Local registration mode	List order			Random order		
	Reader A	Reader B	Two readers	Reader A	Reader B	Two readers
Padding mode	57/60 SE = 0.03	56/60 SE = 0.03	113/120 (94.2%)	57/60 SE = 0.03	57/60 SE = 0.03	114/120 (95.0%)
Non-Padding mode	60/60 SE = 0	56/60 SE = 0.03	116/120 (96.7%)	60/60 SE = 0	59/60 SE = 0.02	119/120 (99.2%)
pool both modes	117/120 (97.5%)	112/120 (93.3%)	229/240 (95.4%)	117/120 (97.5%)	116/120 (96.7%)	233/240 (97.1%)

Reader agreement

We define reader agreement as the percentage of FOVs for which the two readers obtained a similar registration error (smaller or greater than $5 \mu\text{m}$) on the same FOV. For registration before focusing the microscope, the two readers obtained 77% agreement. For registration after focusing (both modes), the agreement between the two readers was higher than 95%. Figure 6 shows one outlier from registration after focusing the microscope. Figure 6-B and 6-C show the two camera images are focused differently by the two readers. The clearer image (6-B) obtained a registration error less than $5 \mu\text{m}$, while the other one (6-C) was approximately $400 \mu\text{m}$. We believe that this is related to the visual acuity of the two readers. To avoid this problem the microscope should be tuned to each reader during global registration. First, the reader should adjust the eyepiece so the reticle is in focus. Next, the reader should focus the microscope on some feature in the FOV. Finally, the system administrator should adjust the camera so the camera image of the specimen is in focus.

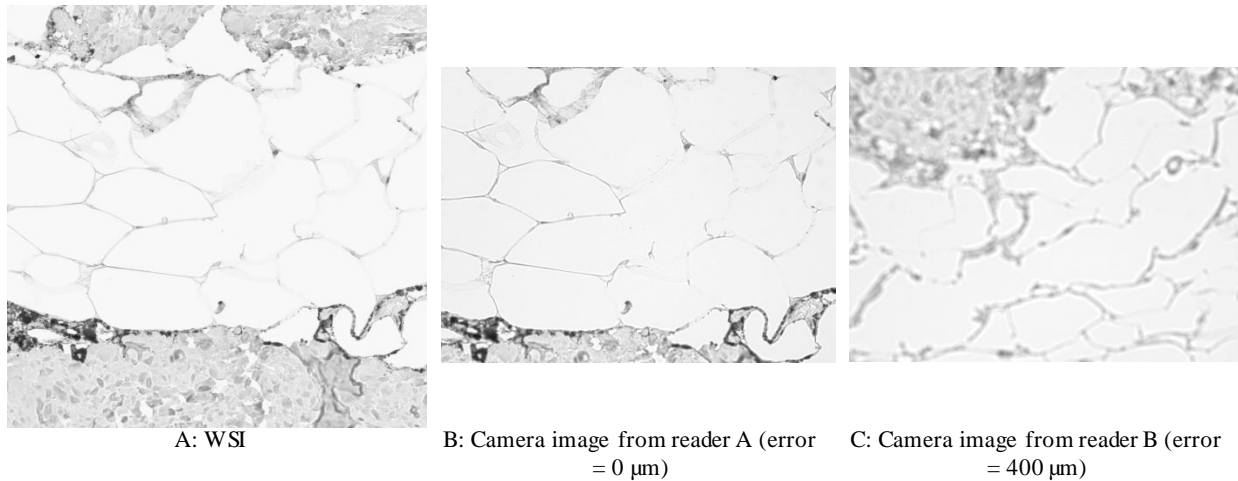


Figure 6: Outlier FOV where two readers obtain disagreement

4. DISCUSSION AND CONCLUSION

eeDAP, an Evaluation Environment of Digital and Analog Pathology, is a tool for conducting studies in which pathologists can evaluate the same FOVs, cells, or features in a glass slide on a microscope and in a WSI on a digital display. Registration accuracy between the microscope and WSI FOVs is a critical feature of eeDAP.

This work shows that the registration accuracy of eeDAP is within $5 \mu\text{m}$ when local registration methods are used after focusing a FOV at 20X and 40X. This accuracy is as small as we could measure with our reticle; the smallest division is $2.5 \mu\text{m}$ at 40X and $5.0 \mu\text{m}$ at 20X. The accuracy does not appear to be affected by tissue types, stains,

scanning magnification, distance between FOVs, and readers. The main factors affecting registration accuracy are the microscope focus quality, scan quality, and FOV content, such as contrast and homogeneity. Future work should investigate these issues and improve the registration methods.

We recommend a pre-study examination of the registration accuracy of any study to be conducted prior to a pivotal study. This examination can help the study designer to decide whether local registrations are effective before and after focusing the microscope or whether FOVs should be changed before starting the study.

ACKNOWLEDGMENTS

Qi Gong and Benjamin P. Berman are supported by an appointment to the Research Participation Program at the Center for Devices and Radiological Health administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration. The mention of commercial entities, or commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such entities or products by the Department of Health and Human Services or the U.S. Food and Drug Administration.

REFERENCES

- [1] B. D. Gallas, M. A. Gavrielides, C. M. Conway *et al.*, "Evaluation environment for digital and analog pathology: a platform for validation studies," *J Med Imaging (Bellingham)*, 1(3), 037501 (2014).
- [2] M. A. Gavrielides, C. Conway, N. O'Flaherty *et al.*, "Observer performance in the use of digital and optical microscopy for the interpretation of tissue-based biomarkers," *Analytical Cellular Pathology*, 2014, 157308 (2014).
- [3] L. Pantanowitz, P. N. Valenstein, A. J. Evans *et al.*, "Review of the current state of whole slide imaging in pathology," *J Pathol Inform*, 2, 36 (2011).
- [4] R. S. Weinstein, A. R. Graham, L. C. Richter *et al.*, "Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future," *Hum Pathol*, 40(8), 1057-69 (2009).
- [5] L. Pantanowitz, J. H. Sinard, W. H. Henricks *et al.*, "Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center," *Arch Pathol Lab Med*, 137(12), 1710-22 (2013).
- [6] L. J. P, "Fast Normalized Cross-Correlation," *Vision interface*, 10(1), 120-123 (1995).
- [7] O. Sertel, J. Kong, H. Shimada *et al.*, "Computer-aided Prognosis of Neuroblastoma on Whole-slide Images: Classification of Stromal Development," *Pattern Recognit*, 42(6), 1093-1103 (2009).
- [8] R. Bala, and R. Eschbach, "Spatial color-to-grayscale transform preserving chrominance edge information," *12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*, 82-86 (2004).
- [9] P. A. Bautista, and Y. Yagi, "Staining correction in digital pathology by utilizing a dye amount table," *J Digit Imaging*, 28(3), 283-94 (2015).
- [10] T. L. Sellaro, R. Filkins, C. Hoffman *et al.*, "Relationship between magnification and resolution in digital pathology systems," *J Pathol Inform*, 4, 21 (2013).
- [11] S. Chenier, and M. Dore, "Oral malignant melanoma with osteoid formation in a dog," *Vet Pathol*, 36(1), 74-6 (1999).
- [12] L. A. Doyle, W. Yang, L. V. Abruzzo *et al.*, "A multidrug resistance transporter from human MCF-7 breast cancer cells," *Proc Natl Acad Sci U S A*, 95(26), 15665-70 (1998).